Development of a Standardized Behavioral Assessment of Therapist Competence in Cognitive Therapy for Depression

Research Thesis

Presented in partial fulfillment of the requirements for graduation

with research distinction in Psychology in the undergraduate

colleges of The Ohio State University

by

Iony D. Schmidt

The Ohio State University

April 2015

Project Advisor: Dr. Daniel R. Strunk, Department of Psychology

Additional committee member: Dr. Amelia Aldao, Department of Psychology

Abstract

The assessment of therapist competence in providing cognitive therapy (CT) is vital to both research and clinical work. In studies examining CT, being able to characterize the competence of therapists is important for placing research findings in context. In clinical settings, therapist competence evaluations can be used for credentialing or selecting therapists to provide clinical services. Currently, the most widely used method of assessing therapist competence is the Cognitive Therapy Scale (CTS; Young & Beck, 1980), an observer-rated measure completed on the basis of CT session recordings. While studies have shown utility of the CTS, a standardized assessment may have important advantages. In this study, I used data from a clinical trial (DeRubeis et al., 2005) to analyze the relationship between a number of patient variables assessed prior to treatment (i.e., history of illness, demographics and life circumstances, family history of mental illness, cognitive dysfunction, functioning, depressive symptoms, and personality disorder status) and CTS scores. While these patient variables failed to predict CTS scores, there was a non-significant trend suggesting observer ratings of another patient characteristic (i.e., patient difficulty) may be related to lower CTS scores. Two patient variables emerged as predictors of patient difficulty ratings: dysthymic disorder and personality disorder. In addition, I developed a standardized behavioral measure of therapist competence in CT for depression. The new assessment uses a series of hypothetical scenarios for which the respondent is asked to role play the cognitive therapist with the assessor acting as the patient, according to a manual detailing how the patient will respond in each scenario. Role plays are being recorded and evaluated using an observer-rated competence scale. Ultimately, I plan to evaluate the reliability and validity of this standardized measure and compare it to the CTS ratings.

Keywords: competence, cognitive therapy, depression

Development of a Standardized Behavioral Assessment of Therapist Competence in Cognitive Therapy for Depression

Major depressive disorder is the single most common psychological disorder, with an estimated lifetime risk of 17%, with point prevalence among adults being between 2 and 4% (Kessler, 2014). Fortunately, a number of treatment options are available. Cognitive therapy (CT), a form of psychotherapy where the therapist helps the patient in identifying and changing negative ways of thinking, has been shown to be efficacious in the treatment of depression (Strunk & DeRubeis, 2001; DeRubeis, Webb, Tang, & Beck, 2010). Over several decades of testing, the CT model and its treatment application have continued to gain empirical support (Beck, 2005). However, there is substantial variability in treatment response and existing research has yet to provide a compelling account of whether and how therapist behaviors play a role in the efficacy of CT (Webb, DeRubeis, & Barber, 2010).

One particular variable of focus is therapist competence; this refers to the skillfulness of the therapist in providing treatment consistent with the respective treatment manual's objectives and guidelines (APA Presidential Task Force, 2006; Waltz et al., 1993). In an analysis of moderate to severely depressed patients, competence ratings were found to predict session-to-session symptom change across the early sessions of CT (Strunk et al., 2010). These effects were particularly strong among patients with a chronic form of depression, higher anxiety, and an early age of first depression onset. Exploratory analysis of the individual CTS items suggested that setting and following an agenda was the item that most strongly predicted subsequent symptom change. Nonetheless, this finding should be interpreted with caution as individual items scores exhibit considerably lower reliability than do total scores (Dobson, Shaw, & Vallis, 1985; Williams, Moorey, & Cobb, 1991).

The assessment of therapist competence in providing CT is vital to both research and clinical work. In research studies examining CT, being able to characterize the competence of therapists is important for placing research findings in context, such as determining empirically whether differences in competence level of therapy provided is associated with differences in symptom relief. In clinical settings, therapist competence evaluations can be used for credentialing or selecting therapists to provide clinical services (Academy of Cognitive Therapy, n.d.).

Current Competence Measure: Cognitive Therapy Scale

Currently, the most widely used method of assessing therapist competence is the Cognitive Therapy Scale (CTS; Young & Beck, 1980). An observer of a CT session can complete the 11-item scale to provide an index of how competently a therapist is providing CT. The scale is divided into two rationally-defined subscales, General Therapeutic Skills (including items such as "Agenda" and "Feedback") and Conceptualization, Strategy, and Technique (including items such as "Guided Discovery" and "Focusing on Key Cognitions or Behaviors"). However, a total score is most often used and is supported by factor analyses that show the CTS to be captured by one factor (Vallis, Shaw, & Dobson, 1986). Each item is rated on a 0 to 6 point scale with total scores ranging from 0 to 66. By convention, scores 40 and above are considered to reflect adequate competence (Young & Beck). However, this cut-off was suggested without any empirical justification.

Several studies have reported evidence for the reliability and validity of the CTS. For example, Vallis and colleagues (1986) reported intra-class correlation coefficients for the CTS of .59 for a single rater and .77 when adjusted for the use of two raters. Strunk and colleagues (2010) reported the same score of .77 when adjusted for two raters. Moreover, several studies

have found that competence ratings of a therapist in working with a specific patient are positively associated with those patients' outcomes. However, these associations have largely involved the relation of competence ratings from some time during CT with therapeutic outcomes achieved across the full course of treatment. Thus, the relation identified could reflect higher competence scores being more common among patients who have already shown a good initial response to CT. As noted previously, one study has found competence ratings predicted subsequent therapeutic gains (Strunk et al., 2010). Thus, this provides some evidence for the predictive validity of the CTS. Nonetheless, as I detail below, there are several important limitations to using the CTS as a measure of therapist competence.

Problems with the CTS

First, in some studies, the CTS has exhibited very poor inter-rater reliability, with correlations below .1 among expert raters (Jacobson & Gortner, 2000). Correlations this low show should be cause for serious concern and call attention to the need for better rater training or a more reliable assessment. This estimate is obviously inconsistent with the higher reliabilities reported by Vallis and colleagues (1986) and Strunk and colleagues (2010). It is difficult to know with certainty what accounts for this discrepancy in reliability estimates. However, one plausible explanation is that these differences are attributable to the extent to which raters are trained together. The raters who contributed scores in Jacobson and Gortner's report had not trained together. Raters in the two studies yielding higher reliability estimates had trained together in the use of the CTS. Also consistent with this possibility, at a Cognitive Therapy Competency Conference held in 1998, experts who largely had not trained together met and rated several recorded sessions of CT using the CTS. These ratings exhibited alarmingly low levels of agreement (Robert J. DeRubeis, personal communication, 2002).

If rater training is key to ensuring adequate reliability of CTS scores, it is important to understand what about this training facilitates agreement. Some CTS items ask raters to evaluate higher level constructs (e.g., the quality of a therapist's strategy for change), leaving it to each individual rater to make important judgments about how specific therapist behaviors relate to this construct. Even the more behaviorally specific CTS items are meant to encompass the therapist's competence in using multiple strategies. For example, the agenda item assesses not only how well the therapist sets an agenda, but how well he or she worked with the patient to follow the agenda. Therefore, raters are left to determine how much to weigh specific therapist behaviors in making their ratings. Perhaps raters who train together in the use of the CTS come to share a common understanding of what behaviors they will consider in rating each item, and which behaviors are most important for each item rating.

If raters must work together to share a common understanding in order to generate reliable CTS scores, it remains unclear whether distinct rater groups would come to the same shared understanding. If they do not, seemingly reliable CTS scores with raters who trained together could still yield poor reliability when assessed across raters who had not trained together. In a study conducted by Strunk, DeRubeis, and Conklin (2015), CT experts were asked to indicate which of 103 specific therapist behaviors were relevant to rating each of the 11 CTS items. There was only moderate agreement among the experts (kappa = .42). This suggests that individual differences in which therapist behaviors are deemed relevant to CTS items may undermine the reliability of the CTS. Thus, an inadequately specific definition of competence may contribute to the poor reliability of competence scores, when examined among experts who have not trained together in use of the instrument.

Second, because the CTS is evaluated in the context of therapists' work with specific patients, these ratings often have the potential to be confounded with the characteristics of the patients included in the rated sessions. A number of patient characteristics have been shown to predict outcome in CT for depression. For example, with respect to personality traits, higher levels of neuroticism have been linked to less symptom change (Klein et al., 2011; Quilty et al., 2008) while extraversion, agreeableness, and likability have been shown to be related to greater symptom change (Bagby et al., 2008, Sasso & Strunk, 2013, Quilty et al., 2008). In the trial from which data for this study are drawn, Fournier and colleagues (2009) found that chronic depression, older age, and lower intelligence predicted poorer treatment response across CT and medication conditions. Being married, unemployed, and having experienced a greater number of recent stressful life events predicted greater response to CT relative to medication. In another paper examining personality disorders as a predictor of response in the same sample, Fournier and colleagues (2008) found that patients who met criteria for a personality disorder showed less symptom change across both CT and medication conditions. In addition, patients with a personality disorder also fared less well in CT than in the medication condition.

The developers of the CTS have assumed that raters will be able to appropriately adjust for patient characteristics. This may or may not be true. For example, a barely adequate therapist who is working with a very easy, compliant patient could be misjudged to have satisfactory competence. Alternatively, an excellent therapist could be misjudged to be mediocre when working with a very difficult patient. That is, it is unclear if raters are able to correctly adjust for the impact of individual differences among patients as they evaluate a therapist's competence.

If a therapist is evaluated using recordings of patient who are either quite compliant or especially difficult, the sampling of patients may bias raters' estimates of that therapist's

competence. Using a small number of observations to assess therapist competence with the CTS assumes that a therapist's competence with a few cases will generalize to a much larger number of cases. For example, it could be that therapists truly exhibit greater competence with less difficult patients. If CTS ratings are only made for less difficult patients, an overestimate of the therapist's competence may result.

This concern is not a trivial issue. Convenience samples of therapy sessions are often part of high stake evaluations of therapist competence. Therapists are evaluated on the basis of cases they select from their own caseloads, allowing for the possibility that these selections could introduce bias in competence evaluations. This reliance on the self-selection of a therapy session for evaluation is utilized by the Academy of Cognitive Therapy, the primary certifiers of CT competence (Academy of Cognitive Therapy, n.d.). To get a representative competence score, a therapist might be asked to pick a session at random, but this may be both practically difficult as well as being difficult to verify. Even if these obstacles could be overcome, therapists may work with different populations of patients, invariably leading to any sessions they would submit for evaluation reflecting not only their own competence, but potentially the characteristics of the population with whom they work.

To understand the role of patient characteristics and therapist competence in patients achieving successful therapeutic outcomes, it is important that measures of patient characteristics and therapist competence accurately reflect the constructs they are intended to measure. In a recent conceptual analysis of psychotherapy process research, DeRubeis and colleagues (2013) suggested that psychotherapy process variables such as therapist competence are likely differentially related to therapeutic outcomes across patients. They characterize patients as falling along a continuum of patient response patterns, with patients ranging from spontaneously

remitting to intractable. Whereas a spontaneously remitting patient is one who would recover with little to no therapeutic intervention, an intractable patient is one who will not get better even if the most potent intervention is provided. Assuming that this range of patients exists in samples of patients with depression, DeRubeis and colleagues suggested that the strongest relation of competence and outcome would be expected among the pliant patients, those who are capable of responding, but only if a high quality treatment is provided. In the absence of reasonably pure measures of competence (i.e., not contaminated by patient characteristics), understanding such effects will likely be more difficult.

Advantages of a Standardized Assessment

While little is known about the degree to which patient characteristics or therapist's process of selecting sessions for evaluation might bias competence ratings, there are reasons to believe that a standardized assessment of therapist competence would have important advantages (American Board of Professional Psychology, 2015). Such an assessment device might be constructed to allow for greater reliability, less potential contamination with patient characteristics, and no vulnerability to biases potentially introduced by therapists selecting work samples for evaluation.

Utilizing standardized patients may reduce the error that is currently accompanied with use of the CTS. In Imel and colleagues' (2014) motivational interviewing study evaluating therapist adherence, performance of therapists with standardized and real patients was compared. Across five measures of adherence, the magnitude of the correlations varied widely, but was as low as an *r* of .04. Their analyses also suggested that it would take fewer standardized patient assessments than real patient assessments to achieve a reliable estimate of adherence. The use of standardized patients therefore shows promise in evaluating therapist behaviors, potentially

including the assessment of therapist competence. By using standardized patients, a competence assessment would not be vulnerable to the potential confounds of patient characteristics and therapeutic gains.

Development of a Standardized Assessment

To counteract the pitfalls of the CTS in measuring therapist competence, I developed a standardized behavioral assessment of therapist competence in CT for depression. This assessment consists of a series of role plays where the therapist participating in the evaluation plays the role of the cognitive therapist and the person administering the assessment plays the role of the patient. There are a total of seven scenarios where the assessor reads the scenario description, including information about the patient and the work that the therapist and patient will be focusing on in the scenario. Each role play lasts 5-12 minutes, with the total administration time ranging from 75 to 90 minutes. A manual is provided for the assessor to follow; this manual contains scenario prompts that are read to the respondent as well as specific guidelines on how the patient is to respond in each scenario. Each scenario is rated with three 0 to 6 point Likert scale items regarding interpersonal skills, technical aspects and competence as a cognitive therapist overall.

This measure was developed with the intention of both assessing the same domains reflected in the original CTS and aspects of competence that experts identified as most essential to therapist competence (Strunk, DeRubeis, & Conklin, 2015). The seven role play scenarios include two involving structuring sessions, two involving behavioral strategy usage, and three involving cognitive strategy usage. These scenarios emphasize important skills judged by experts to be important to the evaluation of therapist competence. In addition, the scenarios focus primarily on intervention and situations that occur in early to mid-treatment. Future efforts to

develop scenarios to assess later sessions or other therapeutic skills could be revisited at a later time.

Role-plays with specific patients and session goals will allow for a standardization of the context in which therapist competence is judged. For example, one of the role-plays has the assessor introduce a 40 year-old patient who is beginning their second session of CT; after providing a description of the patient's goals for CT, the assessor then asks the therapist to work with the patient to set an agenda for their current session. Across the scenarios included, the manual calls for the assessor to vary the description of patient characteristics (including depression severity, demographics, and history of illness) and responsiveness to therapist questioning. Critically, these same scenarios and role-played patients can be used with any therapists who complete the assessment. Thus, the seven scenarios span a diversity of patients, and could help reduce the error that may stem from the patient-related confounding factors that could be a concern with using the CTS.

Study Objectives

I had two specific goals for this thesis. First, I planned to examine the relationship between patient characteristics and CTS scores of the therapists in the DeRubeis et al. (2005) study. My plan was to examine two types of patient characteristics. The first type was pre-treatment characteristics. These characteristics included all of those examined by Fournier and colleagues (2009) as well as an intake depressive symptom severity and the presence of comorbid personality disorders. The second type of patient characteristic is interpersonal difficulty, as observed in the first session of CT. I predicted that patient characteristics, both those from the intake assessment (particularly those previously shown to predict outcome) and session 1 ratings of patient difficulty, would predict the CTS ratings of the therapists in the trial.

Second, an ongoing part of this project involves psychometrically evaluating a newly developed standardized behavioral assessment of therapist competence. At this time, I am in the recruitment and data collection phase of this part of the study. Although this work remains ongoing, I predict that the standardized assessment and coding system will yield high reliability and provide a more valid assessment of competence. With regard to validity, I predict that the standardized competence assessment will be more strongly associated with the degree of cognitive therapist expertise, when compared to original CTS competence ratings**.**

Methods

Sample 1

Participants.

Patients. A total of 60 patients, aged 18 to 70 years, with moderate to severe major depressive disorder, who participated in the CT condition of DeRubeis and colleagues' (2005) clinical trial, were used for this study. These patients were randomly assigned to participate in 16 weeks of CT, which was provided at two sites: the University of Pennsylvania in Philadelphia, Pennsylvania and Vanderbilt University in Nashville, Tennessee.

Exclusion criteria included: (1) history of bipolar disorder; (2) substance abuse or dependence needing treatment; (3) current or past psychosis; (4) another DSM-IV Axis I disorder needing treatment; (5) one of three DSM-IV Axis II disorders (viz., antisocial, borderline, or schizotypal), (6) suicide risk needing hospitalization; (7) medical condition that contraindicated study's medications; and (8) nonresponse to trial of paroxetine.

Therapists. CT was provided by six therapists, four men and two women. Five of the therapists had their PhD while one therapist was a psychiatric nurse practitioner. The therapists each had 5 to 21 years of therapy training, though four of the therapists had 7 to 21 years of CT

training while two of the therapists at Vanderbilt only had 2 years of CT training. The two therapists that had only 2 years of training underwent additional training through the Beck Institute for Cognitive Therapy during the DeRubeis et al., 2005 trial. All therapists followed the standard protocols for providing CT. CT was provided twice per week during the first four weeks of treatment, once or twice per week for the next eight weeks of treatment, and once per week for the last four weeks of treatment.

Measures.

Competence. The CTS is an 11-item observer-rated scale used to assess the competence of a therapist in a therapy session (Young & Beck, 1980). The 11 items include Agenda, Feedback, Understanding, Interpersonal Effectiveness, Collaboration, Pacing and Efficient Use of Time, Guided Discovery, Focusing on Key Cognitions or Behaviors, Strategy for Change, Application of Cognitive-Behavioral Techniques, and Homework. The items are rated on 0 to 6 point Likert-type scales, where higher scores indicate greater levels of competence. Overall scores can range from 0 to 66, where a score of 40 or higher indicates competence.

Two raters (Daniel R. Strunk and Melissa A. Brotman) rated the first four sessions plus week 12 session of CT for each patient. At the time of rating, raters were advanced graduate students who had completed a one year practicum in CT. Sessions were rated sequentially. After reviewing each session, each rater recorded his or her ratings. Then, raters discussed any discrepancies in competence items and agreed upon a consensus rating. These consensus ratings, which are believed to be more valid than independent ratings, were used for all analyses. The ICC for total CTS scores was .77, after being adjusted for two raters.

Patient Difficulty. A single item that asked “How difficult did you feel this client was to work with?” was used to assess patient difficulty. This item was rated on a 0 to 6 point scale,

where 0 reflected that the patient was "not difficult, very receptive" and 6 reflected that the patient was "extremely difficult."

The same two raters who evaluated competence rated this item at the conclusion of session 1 for each patient. The average of these judgments was used in primary analyses. The ICC for patient difficulty ratings was .86, after being adjusted for two raters.

Patient Characteristics. I looked at 41 different patient characteristics as predictors in a series of seven models. These models included the following types of variables: (1) history of illness, (2) demographics and life circumstances, (3) family history of mental illness, (4) cognitive dysfunction, (5) baseline functioning, (6) baseline depressive symptoms, and (7) personality disorder status. Models 1 through 5 are derived from the potential predictors of CT response investigated in Fournier et al.'s 2009 study, while models 6 and 7 were added based on data collected during DeRubeis et al.'s 2005 trial.

History of illness. This model included seven variables: number of prior episodes, onset age, presence of chronic depression, dysthymia (assessed with the SCID-I interview), recurrent depression, atypical depression, and melancholic depression. These variables were assessed through self-report or with the Structured Clinical Interview for DSM-III-R (SCID-I; First, Spitzer, Gibbon, & Williams, 1990).

Demographics and life circumstances. This model included nine variables: age, employment status, gender, marital status, race, number of years of education, income, total number of life events (assessed with the 102-item Psychiatric Epidemiology Research Interview Life Events scale; Dohrenwend, Krasnoff, Askenasy, & Dohrenwend, 1978), and intelligence (assessed with 30-item Shipley–Harford Living Scale; Shipley, 1940)

Family history of mental illness. This model included five variables that reflected the incidence among the patients' first-degree relatives of: (1) major depressive disorder; (2) any other mental disorder; (3) hospitalized for psychiatric reasons; (4) prescribed psychiatric medications; and (5) attempted suicide. These variables were assessed using the Family History-Research Diagnostic Criteria (Andreasen, Endicott, Spitzer, & Winokur, 1977), modified to yield the diagnoses in the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM–IV; American Psychiatric Association, 1994).

Cognitive dysfunction. This model included five variables: attributional style (assessed with the 12-vignette, self-report Attributional Styles Questionnaire; Seligman, Abramson, Semmel, & von Baeyer, 1979), perfectionism and need for approval (each assessed with the 40-item self-report Dysfunctional Attitude Scale; Weissman & Beck, 1978), self-esteem (assessed with the 10-item Rosenberg Self-Esteem Scale; Rosenberg, 1965), and hopelessness (assessed with the 20-item, self-report Hopelessness Scale; Beck, Weissman, Lester, & Trexler, 1974).

Baseline functioning. This model included the following eleven variables: self-report anxiety (assessed with the Beck Anxiety Inventory; Beck, Epstein, Brown, & Steer, 1988), interviewer-evaluated anxiety (assessed with the Hamilton Rating Scale for Anxiety; Hamilton, 1959), anxiety sensitivity (assessed with the 16-item self-report Anxiety Sensitivity Index; Reiss, Peterson, Gursky, & McNally, 1986), global assessment of functioning (American Psychiatric Association, 2000), presence of any axis I comorbidity (assessed with the SCID-I), positive affect and negative affect (assessed with the Positive and Negative Affect Scale; PANAS; Watson, Clark, & Tellegen, 1988), and the five traits of Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness from the Five-Factor model of personality (assessed with the 60-item NEO-Five Factor Inventory; Costa & McCrae, 1992).

Baseline depressive symptoms. Depressive symptoms were assessed with two measures: the Beck Depression Inventory-II (BDI; Beck et al., 1996) and the 17-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960), modified to include atypical symptoms (DeRubeis et al., 2005). The BDI is a 21-item self-report measure, where individual items are scored 0 to 3 and total scores can range from 0 to 63. Total scores on the 17 item HRSD range from 0 to 52. The measure used as the primary indicator of session-to-session depression severity was the BDI.

Personality disorder status. Personality disorder status was assessed with the Structured Clinical Interview for DSM-III-R Personality Disorders (First, Spitzer, Gibbon, & Williams, 1990). As noted above, patients who met criteria for antisocial, schizotypal, and borderline personal disorders were excluded from the study.

Analytic Strategy. I used linear regression and correlation analyses to investigate the relation patient characteristics as predictors of CTS scores, patient difficulty as a predictor of CTS scores, and patient characteristics as predictors of patient difficulty.

For analyses in which I examine patient characteristics as predictors of other variables, I followed an approach used by Fournier et al. (2009) in their analysis of patient characteristics as predictors of treatment response in the DeRubeis et al. (2005) trial. Specifically, I planned to test a series of models examining sets of conceptually related predictor variables (i.e., those listed above under Patient Characteristics). For each of these models, I first planned to evaluate the statistical significance of the overall model. For any model that was significant, I would retain only the variables within the models that had a significance value of $p < .20$ and re-evaluate the model. Then, I would retain only the variables that had a significance value of $p < .10$ and again

retest the models. Finally, I would retain the variables from the last step that had a significance value of $p < .05$.

In analyses where I looked at CTS scores as predictors of session-to-session symptom change, I used an approach where I took CTS scores from sessions 2, 3, and 4 and controlled for CTS scores at sessions 1, 2, and 3. This analysis was a repeated-measures regression analysis, implemented using SAS PROC Mixed (without specification of random effects). In this analysis, a vector of lagged BDI scores for each participant (i.e., BDI scores from sessions 2 through 4) served as the dependent variable, with BDI scores from the previous sessions entered as a covariate (i.e., BDI score at Session 1 serves as a covariate in predicting BDI score at Session 2, etc.). A variable reflecting CTS scores at sessions 2, 3, and 4 was examined as an additional predictor. Thus, these models use repeated observations to estimate the association between CTS scores (across Sessions 2–4) and BDI scores in those same sessions while controlling for BDI scores in the previous session. In these analyses, a negative *t* score would indicate that higher CTS scores are associated with greater symptom prior symptom change (from the previous to the current session). On the basis of a comparison with alternative covariance structures (viz., compound symmetry, Toeplitz, and first-order autoregressive), unstructured was identified as achieving the best fit for this model (on the basis of Akaike's information criterion, Schwarz's Bayesian criterion, and -2 res log likelihood).

Sample 2: Ongoing Study

Participants. Participants include therapists who have been trained in CT for depression and one of the patients they are currently treating. Training in cognitive therapy can include any experience with cognitive therapy, whether it be from reading a book on CT, attending a CT workshop, or actually practicing CT. The total number of participants is dependent on the

number of cognitive therapists and their patients who are willing to participate in the proposed study. Therapists are being recruited through email requests and Facebook postings.

Inclusion criteria for participation in the study are that the participating therapist must be: (a) a therapist with experience/training in cognitive therapy and (b) able and willing to give informed consent to participate in this study. They must also have a patient that is: (a) in therapy addressing mood disorders with the participating therapist and (b) able and willing to give permission to be audio recorded for the CTS submission.

Measures.

Demographics. A 5-10 minute survey is conducted online through Qualtrics.com. It includes an electronic consent form and description of the study. If participants agree to give consent, they are given access to the demographics section of the survey. This survey includes questions regarding contact information, age, ethnicity, education, and experience with learning, providing, and supervising CT.

Competence. (View appendix A). As described above for sample 1, we also used the CTS to assess competence in sample 2. Participating therapists are asked to send an audio recording of one of their therapy sessions with a consenting patient to be rated using the CTS. Ratings will be conducted by raters at the graduate and post-graduate level who are familiar with the CT and the CTS rating manual.

CT experience. CT experience will be defined as a continuous variable as assessed through questions concerning the number of years a therapist has with training or providing CT.

Standardized Behavioral Assessment of Therapist Competence in Cognitive Therapy for Depression (SBA). Assessments are administered by a single person, and are conducted over phone calls and audio recorded for rating purposes. Ratings are performed by raters at the

graduate and post-graduate level who are familiar with the developed rating manual. Independent raters will evaluate the CTS and the standardized behavioral assessments.

Procedure. I am recruiting therapists with varying levels of CT training through use of listservs, Facebook posts, and word of mouth. Interested therapists are given a link to an electronic consent form and demographics Qualtrics survey. Following the survey, eligible participating therapists are requested to submit to Box.com an audio recording of a CT session to be rated using the CTS and to participate in the audio recorded standardized behavioral assessment. After completion of the study, participating therapists and their patients involved with the CTS submission will be compensated with Amazon.com gift cards.

Analytic Strategy. I plan to evaluate rater reliability for both the CTS and SBA. I plan to examine the relationship of SBA scores with CTS scores. I also plan examine the magnitude of the relation of CTS and CT years of experience with the magnitude of the relation of SBA and CT years of experience.

Results

Prior to examining the specific tests of interest, I characterized the variables of interest with descriptive statistics (see Table 1).

Patient Characteristics as Predictors of CTS

Using regression, I examined patient characteristics as predictors of the therapists' CTS scores. As described in the analytic strategy section, I examined 7 models, each testing a different set of patient characteristics as predictor variables. As described, only if the overall model were significant did I plan to examine specific predictors. For each of the seven models, the overall model was non-significant (all *p*s > .05), meaning that I failed to find a significant relationship between patient characteristics and therapists' CTS scores.

Patient Difficulty and CTS Relation

Next, I examined the relation between patient difficulty and CTS scores. There was a non-significant trend for patient difficulty ratings made from the observation of the session 1 to be related to average CTS scores from all sessions (r = -.24, p = .07). Although there was no evidence of serious non-normality in either variable, a non-parametric test of this relation (i.e., a Spearman correlation) suggested a significant relation (r_s = -.27, p = .04). This indicates that higher ratings of patient difficulty are negatively associated with therapists' CTS scores.

Exploratory: Patient Characteristics Predicting Difficulty

I next examined patient characteristics as predictors of observer-rated patient difficulty. To do so, I used the same approach above to test the 7 overall models and proceeded to examine individual predictors if warranted.

Of the seven models tested, two yielded significant overall models: (1) history of illness and (2) personality disorder status. Reducing models using the procedure described above, the history of illness model led to the identification of one significant predictor: dysthymic disorder. In the personality disorder model, the single predictor examined, presence of a personality disorder, was significant. To generate a final model, I examined these two predictors together. The resulting overall model including these two patient characteristics as predictors of patient difficulty was significant, R^2 = .21, F(2,59) = 7.51, p = .001. More specifically, the presence of dysthymic disorder (β = -.36, t = -3.07, p = .003) was related to higher CTS scores, while the presences of a personality disorder (β = .27, t = 2.29, p = .02) was related to lower CTS scores.

Exploratory: Patient Characteristics that Predicted Difficulty as Predictors of CTS

In an exploratory analysis, I then examined the patient characteristics that predicted patient difficulty (viz., dysthymic disorder and presence of a personality disorder), as predictors

of CTS scores. However, neither the overall model (R^2 = .06, $F(2, 59) = 1.71$, p = .19) nor either of the individual predictors was significant (for dysthymia disorder: β = .20, $t(59) = 1.56$, p = .13; for personality disorder status β = -.12, $t(59) = -.96$, p = .34).

Prior Symptom Change and CTS Ratings

Prior symptom change and CTS: session-to-session data. I examined a session-to-session model in which CTS ratings at session 2, 3, and 4 were examined as predictors of BDI scores at sessions 2, 3, and 4 controlling for BDI scores at sessions 1, 2, and 3. In this model, CTS scores were not identified as a significant predictor, b = -.03, SE = .03, $t(58) = -1.05$, p = .30.

Prior change and CTS: Intake through week 12, regressed change. Using CTS scores from the session occurring in the twelfth week of treatment, I tested whether these scores were related to prior symptom change, using a regressed change approach (Cohen, Cohen, West, & Aiken, 2003). To do so, I examined CTS ratings from week 12 as a predictor of HRSD scores at week 12, while controlling for HRSD scores at intake. In this model, week 12 CTS scores failed to emerge as a significant predictor (β = -.05, t = -1.36, p = .18).

Using the same approach for BDI scores rather than HRSD scores, I examined CTS scores at week 12 as a predictor of BDI scores at week 12, with BDI scores at intake included as a covariate. In this model, week 12 CTS scores failed to emerge as a significant predictor (β = -.22, t = -1.52, p = .14). While non-significant, the direction of both effects indicated that CTS scores were numerically related to greater prior changes in depressive symptoms.

Sample 2

The current study of testing the developed standardized behavioral competence assessment is currently in its recruitment and data collection phase, so analyses are yet to be conducted. To date, 10 therapists have agreed to participate and begun the study.

Discussion

The failure to find a relationship between intake patient characteristics and therapist competence scores does not support my hypothesis that patient variables would predict CTS scores. However, ratings of patient difficulty were related to CTS scores at the level of non-significant trend. This relation was significant in a non-parametric test (i.e., a Spearman correlation). Thus, these results largely failed to find a relation between patient characteristics and CTS scores. However, the evidence of a possible relation of CTS scores and patient difficulty suggests additional research is warranted.

In exploratory analyses, I found a significant relation between particular patient characteristics and ratings of patient difficulty. In particular, presence of dysthymic disorder and presence of personality disorder were related to patient difficulty ratings, with patients with dysthymic disorder being rated as less difficult and patients with personality disorder being rated as more difficult. Thus, these findings failed to reveal any direct relationship with intake patient characteristics and CTS scores, but raised other possibilities. For example, one possibility is that intake patient characteristics contribute to patient difficulty, and that patient difficulty in turn has an impact on CTS scores. In any event, we failed to find any simple relations of intake patient characteristics and CTS scores.

Though my findings largely were contrary to my hypothesis, it is important to place these results in context. First, there was some evidence of a relation between patient difficulty and CTS scores. Second, features of the sample from which data were drawn may have limited my

ability to detect an effect. Patients with certain personality disorders were excluded (viz., antisocial, borderline, schizotypal personality disorders). In addition, patients were limited to those with moderate to severe depressive symptoms. These exclusion criteria may have limited our ability to detect any relationship between these factors (and any related patient characteristics) and CTS scores. In addition, other unexamined patient characteristics may be important. These include the extent to which patients perceive the therapy as logical and therapeutic expectations (Callahan et al., 2009, Carter et al., 2011; Lewis et al., 2012; Meyer et al., 2002).

Even if CTS scores were not contaminated by patient characteristics, there are other potential sources of error that could be contributing to CTS scores as they are often used. As noted in the introduction, these include therapist selection and the population of patients with whom therapists are working.

A standardized assessment could avoid these potential causes for concern. A standardized assessment tool that yields consistently high inter-rater reliability and may provide a more accurate measure of individual therapist's competence is worthy of further study.

Limitations

Though I largely failed to find a relationship between patient variables and therapist competence scores, this could be due to several factors. In addition to the exclusion criteria in the trial, these analyses were also limited by examining only six therapists who were judged as sufficiently expert in providing CT to serve as a therapist in the trial. Furthermore, these therapists were all experienced in providing CT and may not provide a representative sample of the range of competence present among all therapists providing CT.

Second, CTS ratings may suffer from a halo effect, where all items for an individual are rated the same according to an overall impression the rater has of the individual with scores being inflated or deflated accordingly (Blackburn & al., 2001). Being prone to the halo effect can lead to homogenous scores ($\alpha = .95$ on the CTS), meaning that scores fail to reflect differentiation between performances on different items (Vallis et al., 1986). Homogenous and inflated scores would likely attenuate any relation of competence with other variables of interest.

It is important to note that the same raters who evaluated patient difficulty at session 1 also generated CTS ratings. Thus, the relation of patient difficulty and CTS scores might be attributable at least in part to shared measurement error. If correlated measurement error contributed strongly to this relation, I would expect a stronger relation when the measures occurred on the same day rather than across several days or weeks. When I looked at the relation of session 1 patient difficulty with CTS scores at each session, this relation was not particularly strong at session 1.

In addition, as has been documented in research on employee selection (Hattrup, 2012), it is possible that competence evaluations are influenced by the raters' bias in evaluating therapists (demographic factors, similarities to therapist, etc.). Neither the CTS nor the standardized assessment solves this problem, but perhaps the standardized assessment could be modified to address this concern in the future. A possible modification to the standardized assessment is to transcribe the assessments; this could help eliminate rater biasing that stems from therapist characteristics such as gender or race.

Future Directions

Although I failed to find a relationship between patient variables and CTS scores, my analyses do not rule out the possibility of this relationship being present in different and more

varied datasets. Nonetheless, concerns about the use of the CTS persist and the development of reliable and valid approaches to evaluating therapist competence remains critical to research efforts and is key to the evaluation of therapists for clinical purposes.

As I mentioned before, my ongoing study on testing the developed standardized assessment is currently in its data collection phase. I hope that this measure will yield more reliable as well as more accurate scores reflecting therapist competence. A standardized measure should help not only in research settings, but it could also help in clinical settings. This new measure could be used as a more refined dissemination tool that would aid in identifying less competent therapists. In turn, these therapists could receive the additional CT training they need so that they can provide the best treatment for their patients and maximize clinical outcomes.

References

Academy of Cognitive Therapy. (n.d.). *Certification in cognitive therapy.* Retrieved from http://www.academyofct.org/.

American Board of Professional Psychology. (2015). Retrieved from http://www.abpp.org/i4a/pages/index.cfm?pageid=3307.American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Andreasen, N. C., Endicott, J., Spitzer, R. L., & Winokur, G. (1977). The family history method using diagnostic criteria: Reliability and validity. Archives of General Psychiatry, 34, 1229–1235.

APA Presidential Task Force on Evidence-Based Practice. (2006). "Evidence Based Practice in Psychology." *American Psychologist, 61(4),* 271-285.

Bagby, R. M., Segal, Z. V., Kennedy, S. H., Quilty, L.C., McBride, C. C., & Costa, J. T. (2008). Personality and differential treatment response in major depression: a randomized controlled trial comparing cognitive-behavioural therapy and pharmacotherapy." *Canadian Journal of Psychiatry, 53,* 361–370.

Barber, J. P., Sharpless, B. A., Klostermann, S., & McCarthy, K. S., (2007). Assessing Intervention Competence and Its Relation to Therapy Outcome: A selected review derived from the outcome literature. *Professional Psychology, Research and Practice, 38,* 493-500.

Beck, A. T. (2005). The current state of cognitive therapy: A 40-year retrospective. *Archives of General Psychiatry, 62,* 953-959. doi:10.1001/archpsyc.62.9.953.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. Journal of Consulting and Clinical Psychology, 56, 893–897.

Beck. A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory (2nd ed.).* San Antonio, TX: The Psychological Corporation.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. Journal of Consulting and Clinical Psychology, 42, 861–865.

Blackburn, I., Milne, D. L., Baker, C., Standart, S., Garland, A., & Reichelt, F. K. (2001). The Revised Cognitive Therapy Scale (CTS-R): Psychometric Properties. *Behavioural and Cognitive Psychotherapy, 29,* 431-446.

Callahan, J.L., Aubuchon-Endsley, N., Borja, S.E., Swift, J.K. (2009). "Pretreatment expectancies and premature termination in a training clinic environment." *Training and Education in Professional Psychology, 3,* 111–120.

Carter, J.D., Luty, S.E., McKenzie, J.M., Mulder, R.T., Frampton, C.M., & Joyce, P.R. (2011). "Patient predictors of response to cognitive behaviour therapy and interpersonal psychotherapy in a randomised clinical trial for depression." *Journal of Affective Disorders, 128,* 252–261.Costa, P. T., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

Cohen, J., Cohen, P., West, S.G., & Aiken, L. S. (2003). Multiple Regression/Correlation Analysis for the Behavioral Sciences. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

DeRubeis, R. J., Hollon S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., O'Reardon, J. P., Lovett, M. L., Gladis, M. M., Brown, L. L., & Gallop, R. (2005). Cognitive Therapy vs Medications in the Treatment of Moderate to Severe Depression. *Archives of General Psychiatry. 62(4),* 409-416. doi:10.1001/archpsyc.62.4.409.

DeRubeis, R. J., Gelfand, L. A., German, R. E., Fournier, J. C., & Forand, N. R. (2013). "Understanding processes of change: How some patients reveal more than others—and some groups of therapists less—about what matters in psychotherapy." *Psychotherapy Research, 24(3),* 419-428. doi:10.1080/10503307.2013.838654

Dimidgian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., Gallop, R., McGlinchey, J. B., Markley, D. K., Gollan, J. K., Atkins, D. C., Dinner, D. L., & Jacobson, N. S. (2006). "Randomized trial of Behavioral Activation, Cognitive Therapy, and Antidepressant Medication in the Acute Treatment of Adults with Major Depression." *Journal of Consulting and Clinical Psychology, 74*, 658-670.

Dohrenwend, B. S., Krasnoff, L., Askenasy, A. R., & Dohrenwend, B. P. (1978). Exemplification of a method for scaling life events: The PERI Life Events Scale. *Journal of Health and Social Behavior*, 19, 205–229.

First M. B., Spitzer R. L., Gibbon M., & Williams J. B. W. (2001). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition With Psychotic Screen (SCID-I/P W/ PSY SCREEN).* New York: Biometrics Research, New York State Psychiatric Institute.

Fournier, J. C., DeRubeis, R. J., Shelton, R. C., Gallop, R., Amsterdam, J. D., & Hollon, S. D. (2008). Antidepressant medications v. cognitive therapy in people with depression with or without personality disorder. *The British Journal of Psychiatry, 192*, 124-129.

Fournier, J. C., DeRubeis, R. J., Shelton, R. C., Hollon, S. D., Amsterdam, J. D., & Gallop, R. (2009). "Prediction of Response to Medication and Cognitive Therapy in the Treatment of Moderate to Severe Depression." *Journal of Consulting and Clinical Psychology, 77(4),* 775-787.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32,* 50-55.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, & Psychiatry, 23,* 56-62.

Hattrup, K. (2012). *Using composite predictors in personnel selection*. In Neal Schmitt (Ed.) *The Oxford Handbook of Personnel Assessment and Selection*. (pp. 297-319). New York, New York: Oxford University.

Imel, Z. E., Baldwin, S. A., Baer, J. S., Hartzler, B., Dunn, C., Rosengren, D. B., & Atkins, D. C. (2014). Evaluating Therapist Adherence in Motivational Interviewing by Comparing Performance with Standardized and Real Patients. *Journal of Consulting and Clinical Psychology, 82,* 472-481.

Jacobson, N. S., Dobson, K. S., Truax, p. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E. T., & Prince, S. E. (1996). *A Component Analysis of Cognitive-Behavioral Treatment for Depression. Journal of Consulting and Clinical Psychology, 64*.

Jacobson, N. S. & Gortner, E. T. (2000). "Can Depression Be De-medicalized in the 21st Century: Scientific Revolutions, Counter-revolutions and the Magnetic Field of Normal Science." *Behaviour Research and Therapy, 30,* 103-17.

Klein, D.N., Kotov, R., & Bufferd, S. J. (2011). "Personality and depression: explanatory models and review of the evidence." *Annual Review of Clinical Psychology, 7,* 269–295.

Lewis, C.C., Simons, A. D., & Kim, H. K. (2012). "The role of early symptom trajectories and pretreatment variables in predicting treatment response to cognitive behavioral therapy." *Journal of Consulting and Clinical Psychology, 80,* 525–534.

Meyer, B., Pilkonis, P.A., Krupnick, J.L., Egan, M.K., Simmens, S.J., & Sotsky, S.M. (2002). "Treatment expectancies, patient alliance, and outcome: further analyses from the National Institute of Mental Health Treatment of Depression Collaborative Research Program." *Journal of Consulting and Clinical Psychology, 70,* 1051–1055.

Reiss, S., Peterson, R. A., Gursky, M., & McNally, R. J. (1986). Anxiety, sensitivity, anxiety frequency, and the prediction of fearfulness. Behaviour Research and Therapy, 24, 1–8.

Rosenberg, M. (1965). Society and the adolescent child. Princeton, NJ: Princeton University Press.

Sasso, K. E., & Strunk, D. R. (2013). "Thin slice ratings of client characteristics in intake assessments: Predicting symptom change and dropout in cognitive therapy for depression." *Behaviour Research and Therapy, 51(8),* 443-450.

Seligman, M. E., Abramson, L. Y., Semmel, A., & von Baeyer, C. (1979). Depressive attributional style. Journal of Abnormal Psychology, 88, 242–247.

Shaw, B. F., Elkin, I., Yamaguchi, J., Olmsted, M., Vallis, T. M., Dobson, K. S., & Watkins, J. T. (1999). Therapist Competence Ratings in Relation to Clinical Outcome in Cognitive Therapy of Depression. *Journal of Consulting and Clinical Psychology, 67,* 837-846.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. Journal of Psychology, 9, 371–377.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1990). *Structured Clinical Interview for DSM–III–R Personality Disorders* (SCID-II; Version 1.0). Washington, DC: American Psychiatric Press.

Strunk, D. R., Brotman, M. A., DeRubeis, R. J., & Hollon, S. D. (2010). Therapist Competence in Cognitive Therapy for Depression: Predicting Subsequent Symptom Change. *Journal of Consulting and Clinical Psychology, 78,* 429-437.

Strunk, D. R., (2014). An Examination of the Cognitive Therapy Scale: A Study of Expert Raters, Manuscript in Progress.

Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986). The Cognitive Therapy Scale: Psychometric Properties. *Journal of Consulting and Clinical Psychology, 54,* 381-385.

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology,* 78, 200-211.

Waltz, J., Addis, M. E., Koerner, K., Jacobson, N. S. (1993). "Testing the integrity of a psychotherapy protocol: assessment of adherence and competence." *Journal of Consulting and Clinical Psychology, 61(4).* 620-630.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54, 1063–1070.

Weissman, A. N., & Beck, A. T. (1978). Development and validation of the Dysfunctional Attitude Scale: A preliminary investigation. Paper presented at the meeting of the American Educational Research Association, Toronto, Ontario, Canada.

Young, J. & Beck, A. (1980). Cognitive therapy scale: rating manual. Unpublished Manuscript.

Young, J. & Beck, A. (1988). Cognitive therapy scale. Unpublished Manuscript.

Table 1

	M	SD	%
Patient Characteristics: History of Illness			
Number of prior episodes	2.3	2.0	
Age of onset	24.2	12.9	
Chronic depression			45
Dysthymia			23
Recurrent depression			75
Atypical depression			28
Melancholic depression			15
Patient Characteristics: Demographics and Life Circumstances			
Age	40.3	11.5	
Unemployed			227
Female			58
Married/cohabiting			30
Caucasian			78
Number of years of education	14.6	2.5	
Income in thousands of US $	30.7	30.8	
Total number of life events	6.8	4.4	
Intelligence	109.0	10.0	
Patient Characteristics: Family History of Mental Illness (among 1st degree relatives)			
Major depressive disorder	0.3	0.2	
Any other mental disorder	0.3	0.3	
Hospitalized for psychiatric reasons	0.1	0.1	
Prescribed psychiatric medications	0.1	0.2	
Attempted suicide	0.0	0.1	
Patient Characteristics: Cognitive Dysfunction			
Attributional styles	0.3	3.3	
Perfectionism	55.7	16.8	
Need for approval	45.7	10.6	
Self-esteem	22.7	5.1	
Hopelessness	11.2	5.3	

Patient Characteristics: Baseline Functioning			
Beck anxiety	15.	9.9	
Hamilton anxiety	16.8	6.9	
Anxiety sensitivity	25.4	11.1	
Global assessment of functioning	50.3	5.2	
Positive affect	16.3	5.4	
Negative affect	27.6	7.8	
Axis 1 comorbidity			63
Neuroticism	32.4	7.7	
Extraversion	20.7	6.6	
Openness	28.3	6.8	
Agreeableness	28.4	6.5	
Conscientiousness	26.6	8.8	
Patient Characteristics: Baseline Depressive Symptoms			
Hamilton depression at baseline	24.0	3.4	
Beck depression at baseline	30.8	8.7	
Patient Characteristics: Personality Disorder Status			
Presence of a personality disorder			45
Patient difficulty	1.8	1.2	
CTS scores	39.7	9.2	

Table 1. This is a list of the patient characteristics at intake, patient difficulty, and CTS scores and their descriptive statistics taken from DeRubeis et al.'s 2005 trial.

Appendix A

Cognitive Therapy Rating Scale (CTRS)

Therapist: _____ Patient: _____ Date of Session: _____

Tape ID#: _____ Rater: _____ Date of Rating: _____

Session# _____ ( ) Videotape ( ) Audiotape ( ) Transcript ( ) Live Observation

Directions: For each time, assess the therapist on a scale from 0 to 6, and record the rating on the line next to the item number. Descriptions are provided for even-numbered scale points. If you believe the therapist falls between two of the descriptors, select the intervening odd number (1, 3, 5). For example, if the therapist set a very good agenda but did not establish priorities, assign a rating of a 5 rather than a 4 or 6.

If the descriptions for a given item occasionally do not seem to apply to the session you are rating, feel free to disregard them and use the more general scale below:

0	1	2	3	4	5	6
Poor	Barely Adequate	Mediocre	Satisfactory	Good	Very Good	Excellent

Please do not leave any item blank. For all items, focus on the skill of the therapist, taking into account how difficult the patient seems to be.

Part I. GENERAL THERAPEUTIC SKILLS

___1. AGENDA

0 Therapist did not set agenda.

2 Therapist set agenda that was vague or incomplete.

4 Therapist worked with patient to set a mutually satisfactory agenda that included specific target problems (e.g., anxiety at work, dissatisfaction with marriage.)

6 Therapist worked with patient to set an appropriate agenda with target problems, suitable for the available time. Established priorities and then followed agenda.

___2. FEEDBACK

0 Therapist did not ask for feedback to determine patient's understanding of, or response to, the session.

2 Therapist elicited some feedback from the patient, but did not ask enough questions to be sure the patient understood the therapist's line of reasoning during the session or to ascertain whether the patient was satisfied with the session.

4 Therapist asked enough questions to be sure that the patient understood the therapist's line of reasoning throughout the session and to determine the patient's reactions to the session. The therapist adjusted his/her behavior in response to the feedback, when appropriate.

6 Therapist was especially adept at eliciting and responding to verbal and non-verbal feedback throughout the session (e.g., elicited reactions to session, regularly checked for understanding, helped summarize main points at end of session.

___3. UNDERSTANDING

0 Therapist repeatedly failed to understand what the patient explicitly said and thus consistently missed the point. Poor empathic skills.

2 Therapist was usually able to reflect or rephrase what the patient explicitly said, but repeatedly failed to respond to more subtle communication. Limited ability to listen and empathize.

4 Therapist generally seemed to grasp the patient's "internal reality" as reflected by both what the patient explicitly said and what the patient communicated in more subtle ways. Good ability to listen and empathize.

6 Therapist seemed to understand the patient's "internal reality" thoroughly and was adept at communicating this understanding through appropriate verbal and non-verbal responses to the patient (e.g., the tone of the therapist's response conveyed a sympathetic understanding of the patient's "message"). Excellent listening and empathic skills.

___4. INTERPERSONAL EFFECTIVENESS

0 Therapist had poor interpersonal skills. Seemed hostile, demeaning, or in some other way destructive to the patient.

2 Therapist did not seem destructive, but had significant interpersonal problems. At times, therapist appeared unnecessarily impatient, aloof, insincere or had difficulty conveying confidence and competence.

4 Therapist displayed a satisfactory degree of warmth, concern, confidence, genuineness, and professionalism. No significant interpersonal problems.

6 Therapist displayed optimal levels of warmth, concern, confidence, genuineness, and professionalism, appropriate for this particular patient in this session.

___5. COLLABORATION

0 Therapist did not attempt to set up a collaboration with patient.

2 Therapist attempted to collaborate with patient, but had difficulty either defining a problem that the patient considered important or establishing rapport.

4 Therapist was able to collaborate with patient, focus on a problem that both patient and therapist considered important, and establish rapport.

6 Collaboration seemed excellent; therapist encouraged patient as much as possible to take an active role during the session (e.g., by offering choices) so they could function as a "team".

___6. PACING AND EFFICIENT USE OF TIME

0 Therapist made no attempt to structure therapy time. Session seemed aimless.

2 Session had some direction, but the therapist had significant problems with structuring or pacing (e.g., too little structure, inflexible about structure, too slowly paced, too rapidly paced).

4 Therapist was reasonably successful at using time efficiently. Therapist maintained appropriate control over flow of discussion and pacing.

6 Therapist used time efficiently by tactfully limiting peripheral and unproductive discussion and by pacing the session as rapidly as was appropriate for the patient.

Part II. CONCEPTUALIZATION, STRATEGY, AND TECHNIQUE

___7. GUIDED DISCOVERY

0 Therapist relied primarily on debate, persuasion, or "lecturing." Therapist seemed to be "cross-examining" patient, putting the patient on the defensive, or forcing his/her point of view on the patient.

2 Therapist relied too heavily on persuasion and debate, rather than guided discovery. However, therapist's style was supportive enough that patient did not seem to feel attacked or defensive.

4 Therapist, for the most part, helped patient see new perspectives through guided discovery (e.g., examining evidence, considering alternatives, weighing advantages and disadvantages) rather than through debate. Used questioning appropriately.

6 Therapist was especially adept at using guided discovery during the session to explore problems and help patient draw his/her own conclusions. Achieved an excellent balance between skillful questioning and other modes of intervention.

___8. FOCUSING ON KEY COGNITIONS OR BEHAVIORS

0 Therapist did not attempt to elicit specific thoughts, assumptions, images, meanings, or behaviors.

2 Therapist used appropriate techniques to elicit cognitions or behaviors; however, therapist had difficulty finding a focus or focused on cognitions/behaviors that were irrelevant to the patient's key problems.

4 Therapist focused on specific cognitions or behaviors relevant to the target problem. However, therapist could have focused on more central cognitions or behaviors that offered greater promise for progress.

6 Therapist very skillfully focused on key thoughts, assumptions, behaviors, etc. that were most relevant to the problem area and offered considerable promise for progress.

___9. STRATEGY FOR CHANGE *(Note: For this item, focus on the quality of the therapist's strategy for change, not on how effectively the strategy was implemented or whether change actually occurred.)*

0 Therapist did not select cognitive-behavioral techniques.

2 Therapist selected cognitive-behavioral techniques; however, either the overall strategy for bringing about change seemed vague or did not seem promising in helping the patient

4 Therapist seemed to have a generally coherent strategy for change that showed reasonable promise and incorporated cognitive-behavioral techniques.

6 Therapist followed a consistent strategy for change that seemed very promising and incorporated the most appropriate cognitive-behavioral techniques.

___10. APPLICATION OF COGNITIVE-BEHAVIORAL TECHNIQUES (*Note: For this item, focus on how skillfully the techniques were applied, not on how appropriate they were for the target problem or whether change actually occurred.)*

0 Therapist did not apply any cognitive-behavioral techniques.

2 Therapist used cognitive-behavioral techniques, but there were significant flaws in the way they were applied.

4 Therapist applied cognitive-behavioral techniques with moderate skill.

6 Therapist very skillfully and resourcefully employed cognitive-behavioral techniques.

___11. HOMEWORK

0 Therapist did not attempt to incorporate homework relevant to cognitive therapy.

2 Therapist had significant difficulties incorporating homework (e.g., did not review previous homework, did not explain homework in sufficient detail, assigned inappropriate homework).

4 Therapist reviewed previous homework and assigned "standard" cognitive therapy homework generally relevant to issues dealt with in session. Homework was explained in sufficient detail.

6 Therapist reviewed previous homework and carefully assigned homework drawn from cognitive therapy for the coming week. Assignment seemed "custom tailored" to help patient incorporate new perspectives, test hypotheses, experiment with new behaviors discussed during session, etc.